Typicality and Compositionality: The Logic of Combining Vague Concepts

James A. Hampton and Martin L. Jönsson

City University, London and University of Lund.

Contribution to E. Machery, W. Hinzen & M. Werning (Eds.) *Handbook on Compositionality*. Oxford: Oxford University Press.

Address correspondence to:

James A. Hampton

Dept of Psychology

City University,

Northampton Square

London EC1V OHB, UK

Fax: +44 20 70408581

Phone: +44 20 70408520

hampton@city.ac.uk

The principle of compositionality (C) is usually understood in the following way:

(C) The meaning of a complex expression is completely determined by the meaning of its parts and their mode of combination.

According to (C), the meaning of the expression "John loves spinach", for example, depends on the meanings of the three words that it contains, and the Subject – Verb – Object syntactic structure that relates them, *and nothing else*.

As expressed above, (C) is a statement about the semantics of *natural language*. It can also be framed slightly differently so that it becomes a statement about the content of complex *concepts*. It is this formulation, and reasons for deviating from it, that we will be concerned with in this text. We will review several decades of psychological research on *typicality effects* and *non-logical reasoning* about category membership that suggest that explanations can be given for of a wealth of phenomena if concepts are understood as *prototypes* (broadly construed). But if concepts are understood in this way this leads to their combination being non-compositional (in the sense of (C)). The evidence that we will review suggests that when it comes to the combination of prototypes a principle corresponding to something like (C') rather than (C) seems to characterize the situation more accurately.[1]

(C') The content of a complex concept is completely determined by the contents of its parts and their mode of combination, *together with general knowledge.*

The reviewed evidence thus gives us good reason to think that this way of thinking about

[1] The relative merits of principles (C) and (C') are discussed further in Jönsson (2008).

concepts (that is, that they are prototypes and combine according to (C’)) is the most productive in leading to scientific understanding.[2]

Research into typicality effects sprung from the observation that members of most conceptual categories exhibit differences in how representative they are of their categories (Rosch and Mervis, 1975). Robins are more typical birds than penguins, hammers are more typical tools than knifes, and apples are more typical fruits than coconuts. Typicality effects were found to affect a wide range of psychological tasks such as the learning of novel categories (Posner & Keele, 1968), the speed and accuracy of categorizing items in common semantic categories like fruits or vehicles (Hampton, 1979), and the strength of inductive inferences (Rips, 1975). The explanation for all of these effects was the proposal that concepts are represented mentally as *prototypes*, where a prototype is a structured set of descriptive properties that captures what a cluster of items of some category typically have in common and what differentiates them from other kinds of thing (Hampton, 1995).[3] Early prototype models suggested a representation in terms of feature lists (Hampton, 1979; Rosch, 1975) such as “flies”, “has a beak”, “hatches from an egg”. More recently, evidence about the importance of causal relations amongst such features has led to more powerful representational mechanisms such as frames or schemas being proposed (Barsalou & Hale, 1993). Hampton (2006) argued that we should nonetheless consider such more powerful representations as prototypes, for the simple reason that they continue to exhibit typicality and vagueness, contain rich descriptions of typical properties and yet lack clearly expressible

[2] This conclusion is not reached by *a priori* arguments similar to those that have been presented in favour of why natural languages or thought have to be compositional (e.g. Fodor and Lepore, 1991; 1996), why natural language doesn’t have to be compositional (e.g. Schiffer,1987), why natural languages in fact are not compositional (e.g. Lahav, 1989; Pelletier, 1994; 2000) or why thought is not compositional (Schiffer, 2003). The situation is, rather, that a certain approach to conceptual combination has (experimentally) shown itself to be extraordinarily revealing with respect to matters of typicality and non-logical reasoning and that this approach presupposes a non-compositional understanding of concepts.

[3] For a review or recent work on prototypes see Hampton (2006).

definitions. The notion of a prototype has also been employed with great success in different domains of linguistics, including lexical semantics, morphology, syntax and phonology (Taylor, 2003).

Differences in *typicality* need to be distinguished from differences in *degrees of membership* (Kamp & Partee, 1995; Osherson & Smith, 1997). Both robins and ostriches are indisputably birds, but robins are more *typical* birds than ostriches. Nonetheless it seems that in many conceptual categories that show typicality effects, membership also comes in degrees so that it is indeterminate whether something belongs to a certain category or not. Whether carpets and clocks are furniture, for instance, seems to be indeterminate in this way. In effect semantic categorization often shows *vagueness* (Keefe & Smith, 1997), with statements being neither clearly true nor clearly false.

To explain the twin phenomena of vagueness and typicality and their close correspondence it is most parsimonious to suppose that both arise from an underlying dimension of similarity to a prototype (Hampton, 1998). The closer an item is to the prototype representation of a concept, then the more frequently and quickly it will be judged as falling under that concept, the more typical it will be judged to be, the stronger will be inductive arguments from item to category, and the more similar it will be to other category items.

By invoking prototypes we can also provide explanations of how the typicality structure of complex concepts is related to that of their constituents and why a considerable part of our reasoning appears to violate logical and statistical laws. The next two sections will substantiate these two claims. Section four will address some philosophical arguments that have been advanced against prototype theory in virtue of it being non-compositional. While we concede that prototypes are non-compositional in a strict sense, we maintain that they can perform the same explanatory work that concepts that behave compositionally can. This claim is corroborated by a closing discussion of the productivity and systematicity of our conceptual competence, phenomena which are often held to underlie the conviction that concepts have to be compositional.

II. How to Combine Prototypes

Let us take the simplest case for combining two concepts – the formation of a complex concept that refers to their conjunction, as in when the concepts PET and FISH are combined to form PET FISH. Prima facie one may suppose that this should be a paradigm case for compositional combination – the concept of PET FISH should just be the concept that refers to the class of things that are both pets and fish.[4]

But what of the prototypes of these concepts and the typicality of different items within them? When two concepts are combined into a conjunction, how does the typicality of an instance of the conjunctive category relate to its typicality in each of the conjoined categories? This problem has become known as the "guppy" problem, after Osherson and Smith's seminal article of 1981 in which they explored the difficulties inherent in the process.[5] Given that a guppy is typical to degree *x* as a pet, and typical to degree *y* as a fish, how typical should it be as a pet fish?

There are two ways in which an answer can be derived. The first is to try to find a

[4] There are of course many other ways in which complex concepts can be formed, including the formation of non-intersectional compounds like APARTMENT DOG (Kamp & Partee, 1995; Wisniewski, 1997). For simplicity, we do not consider them here.

[5] A version of Stigler's Law of Eponymy (Stigler, 1980) could apply here, since Storms et al. () discovered that the guppy does not actually show the guppy effect. Plenty of other examples do, however. We will continue to use the guppy example here, since it is the case most often cited.

function that would compute the typicality of an instance for the conjunction directly from its typicality in the constituent parts, by way of, for instance, recruiting *fuzzy set theory* (Zadeh 1965). The second is to form a new prototype from the parts of the complex concept and then to use *that* to generate the typicality rating (Cohen & Murphy, 1984; Hampton, 1987; 1988). Figure 1 illustrates the two approaches, which we have labeled the extensional and the intensional routes.

Osherson & Smith (1981) quickly identified major failings with the extensional route. The difficulty they suggested is that whereas a guppy may be an atypical fish (it is not caught in nets and served up with French fries) and may be an atypical pet (it cannot be stroked and petted), it is yet a very typical pet fish. This intuition is at odds with either of the fuzzy logic rules for forming a conjunction, for which membership in a conjunction can never be greater than membership in either constituent set. (Quite reasonably the statement “p AND q” should never be more true than just “p”).

In fact the typicality of an item in the conjunctive class cannot be derived in any simple way from its typicality in each of the individual classes. Osherson and Smith (1982) made this point very forcefully with a demonstration concerning a figure that was half way between a circle and a square. They pointed out that even though the shape may be considered just as good a circle as it is a square, it is not equally good as a round circle versus a round square, being a much better example of a round square than a round circle. This demonstration clearly implies that a search for compositional rules that will take degree of membership (or typicality) in each constituent and derive a value for degree of membership (or typicality) in the conjunction is doomed. The shape has equal membership in the constituents [square] and [circle], but when the identical additional conjunct [round] is added to each, the results are no longer equal.

The conclusions drawn by Smith and Osherson (1984, see also Osherson & Smith, 1997) were that category membership and category typicality were two very different issues. Category membership should follow logical rules so that a creature should only be a pet fish if it is both a pet and a fish. On the other hand typicality has to be derived by means of the intensional route which takes account of the interaction between the contents of each concept. In the case of the round circle and the round square, the concept of "round circle" reduces to the concept of "circle", whereas "round square" involves integrating the two concepts to generate a representation of something that is part round and part square – and so a good match to the actually presented shape. Similarly, Cohen & Murphy (1984) and Hampton (1987; 1988b) suggested that the reason that a guppy is a better pet fish than it is a pet or a fish is that the two concepts are first integrated into a single composite concept. Some of the features of pets and of fish will not be carried through to the new complex concept, and so it is quite possible for an instance (such as a guppy) to match the complex prototype better than it does either of the original constituents.[6]

While the intensional route provides a promising way to understand typicality, it appears too unconstrained as a general theory of how concepts should be combined. In particular, the intensional route is unable to rule out logical inconsistencies in category membership assignments, such as that "x is A" is false whereas "x is both A and B" is true. Effectively, conjunction via the intensional route does not have to correspond to conjunction as it is normatively defined (i.e. as corresponding to the operation of set intersection).

To illustrate the point, let us first consider the case of the conjunction of two concepts

[6] There might also be information – extensional feedback – available from actual experience with pet fish that influence the representation corresponding to the complex concept. See below.

that are well-defined (for the sake of argument). Neither suffers from vagueness. Suppose that BACHELOR is represented as the conjunction of the features [male, adult, unmarried, human] and DOCTOR is represented as the conjunction of the features [MD qualification, practices medicine]. Then a simple aggregation or pooling of the two sets of features via the intensional route gives us [male, adult, unmarried, human, MD qualification, practices medicine] as the necessary set of features for defining a bachelor doctor. Aggregation gives us the appropriate truth table for conjunction, provided there is no vagueness.[7]

Even if one concept is well-defined but the other is vague, there is still no logical problem for conjunction via the intensional route. A “bald bachelor” will have the features that define bachelor coupled with a vague predicate “is bald”, and one can expect that a person would only be considered a bald bachelor if they were both bald and a bachelor.[8]

The problems arise (and the intensional route diverges from logical norms) when both concepts are subject to vagueness at their borderlines. Consider the concepts “red” and “ripe” as applied to apples. Both redness and ripeness come in degrees, and the decision as to where to draw the line of what counts as red and what counts as ripe is to an extent arbitrary and vague. The extensional route would proceed as follows: (1) a decision is made whether the apple is red, and (2) a decision is made whether the apple is ripe, and (3) a truth table rule is applied to the results of (1) and (2) so that the object is a red ripe apple only if the two decisions are positive. So far, so simple. The extensional route accords with logical norms.

How about the intensional route? First a prototype representation of the conjunctive

[7] As a first pass one can also suppose that the two prototypes simply compose – the bachelor doctor will be interested in dating girls and have illegible handwriting just like other bachelors and other doctors.
[8] Some dimensional adjectives work differently – “tall child” means tall *for a child*. We assume here that “bald for a bachelor” has the same baldness criterion as “bald for a married man”.

concept is formed by aggregating the feature [red] with the feature [ripe]. Any instance is then judged to belong in the conjunction on the basis of its *overall similarity* to this composite representation. If redness and ripeness are true of an apple to differing degrees, then of necessity there will be *compensation* between membership in one constituent category and membership in the other. Consider two red apples, both equally on the verge of ripeness to the point where they are neither clearly ripe nor clearly unripe. Suppose also that they are both clearly red – membership in the class "red apples" is not in doubt – but that apple A is actually *redder* than B. Both are red, but A is more typically red, being closer to the prototype shade of red than is B. That means that A will be a better match to the prototypical red ripe apple than B, and as a result should be more likely to be categorized as such. The degree of redness can compensate for the lack of ripeness. That will mean that an apple may be considered unripe *simpliciter*, but because it is very red it may yet be classified as a red ripe apple. Hence "p" is false, but "p AND q" is true.[9]

Figure 2 illustrates the problem. Redness and ripeness are represented as the two axes, with the category borderlines represented as dotted lines passing through the centre of each scale. A conjunctive concept defined by logical class intersection would therefore have the boundary represented by the bold L-shaped line that encloses the top right quadrant where both concept memberships are positive. However the intensional route will mean that membership in the conjunction is a function of distance from the top right corner (the prototype for the conjunction) and hence the category boundary for the conjunction will be curved with a negative slope.[10]

[9] Alternatively if the threshold for "red ripe apple" is raised so as to exclude these cases, it will necessarily also exclude other cases that have in fact passed the threshold for each separate constituent with the result that "p" and "q" are true, but "p AND q" is false.
[10] Billings and Marcus (1983) review similar compensatory rules in the wider context of decision making.

In the case of feature-based prototypes the same problem can be readily illustrated. Figure 3 illustrates the case of combining two prototypes with “polymorphous” membership rules requiring that any 2 out of 3 features be present in order to belong to a category (Dennis, Hampton & Lea, 1973). The composite prototype is constructed by amalgamating the 6 features, possession of any 4 of which is now sufficient for category membership. Item 1, which has 2 A features and 2 B features is therefore in the conjunctive set. But so is Item 2 which as 3 A features and only 1 B feature. Item 2 is therefore in the category $A \wedge B$, but is not a member of the category B.[11]

Osherson and Smith (1997) took this undesirable feature of intensional combination processes as good reason to reject the notion that category membership is based on similarity to a prototype. Explaining typicality judgments requires an intensional combination mechanism, but category membership judgments need to respect the rules of logic. They were led to the conclusion that category membership could not be based on the notion of typicality in the way that prototype theory would propose. Hence category membership and typicality must be separable aspects of conceptual structure, and must require different theoretical treatments (see also Hampton, 2007; Kamp & Partee, 1995).

III. Non-Logical Reasoning

The story so far suggests that we need to stick to the Extensional route if we are to “safeguard … classical logic, with its limpid and sensible rules for deriving truth-values of many compound statements from the truth-values of their parts” (Osherson & Smith, 1997).

[11] Composite prototype concepts do not always have to show these effects. If it is the case that each concept has a single feature that is alone sufficient to reach criterion, and the sum of other features is insufficient on their own to reach criterion, then the two concepts will appear to be well-defined, and their conjunction will likewise be intersective (Hampton, 1995).

However, a psychological account of concepts need not *of necessity* produce rules for combination that agree with classical logic. Being an empirical theory, we need first to discover what people actually do when asked to classify items into conjunctive concepts with vague boundaries. Research on reasoning (Evans, 1982; Evans & Over, 1996) has found that people's reasoning often falls short of all but the most simple of logical rules. Non-logical reasoning is not random or arbitrary, but follows its own heuristic rules with particular adaptive value (Gigerenzer et al., 1999; Chater & Oaksford, 1999).

In the case of concept conjunctions there is ample evidence that people engage in non-logical behavior. More strikingly, this behavior is exactly in line with the type of effect just described, in which membership in a conjunction is not a simple logical combination of membership in each individual category, but instead reflects average similarity to each constituent.

In a series of papers, the first author has shown that there is a complex, yet very systematic, set of processes by which prototypes are combined (Hampton, 1987, 1988a, 1988b, 1996, 1997a, 1997b). The studies considered people's judgments of both extensional and intensional aspects of concept conjunction – the extensional aspect concerning what people consider to belong in a particular category, and the intensional aspect concerning what properties people consider are generally true of a particular category. In both cases, people were asked to provide data for pairs of constituent categories (for example SPORTS and GAMES) and then for their conjunction (SPORTS THAT ARE ALSO GAMES).[12] Extensive further work on this problem has also been conducted by Gert Storms and his colleagues at the Katholik University of Leuven (e.g. Storms et al., 1996; 1998; 1999; 2005).

---

[12] A relative clause construction was chosen as a way of making the intersective interpretation quite explicit. Storms et al. (1998) however have shown that the precise form of the conjunctive expression has no effect on the extent of the non-logical phenomena observed.

Studies of extensions led to the discovery of three main phenomena. First, as expected from the problem analysis presented in Figure 2, people did not follow the logical conjunctive boundaries corresponding to the Extensional Route. Instead they tended to *overextend* their conjunctive categories. Figure 4 illustrates the phenomenon with data from Hampton (1988b). The vertical axis shows the degree to which people judged that an instance was a member of the Head noun category (the first mentioned concept), and the horizontal axis shows the same measure for the Qualifier noun category (the concept in the relative clause). The contours show the degree to which membership in the conjunction varied as a function of each axis. The boundary of the conjunction is shown by the dashed contour line marked 0.0. It is clear that the conjunction borderline passes close to the origin in the centre of the diagram. An item that was exactly on the borderline for each constituent would also be borderline for the conjunction. As a consequence there are regions in the diagram where items that are not in one or other of the constituents are nevertheless in the conjunction. Analysis of individual membership judgments confirmed this pattern – both using ratings of membership averaged across participants and using frequencies of making a positive judgment. Regression analysis, using an interaction term to capture the curvature, was able to predict 90-95% of the variance in conjunctive membership. The same pattern was found when participants judged all three category memberships, (Hampton, 1988) or when different groups made the judgments (Hampton, 1997a). This overextension effect is nicely captured in a quote from Rebecca Goldstein's excellent novel *The Mind-Body Problem* in which the protagonist writes:

"My intelligence, like my beauty, has always been overpraised, misperceived.

The conjunction favors both conjuncts. I am beautiful for a brainy woman, brainy for a beautiful woman, but objectively speaking, neither beautiful nor brainy"

Rebecca Goldstein: The Mind-Body Problem, 1983

The other two phenomena are quickly described. Figure 4 clearly shows an asymmetry in that the area of overextension is greater in the lower right quadrant than in the top left quadrant. This asymmetry reflected the fact that in regressions predicting conjunctive membership from constituent membership, greater weight was accorded to the qualifier noun, resulting in fewer overextensions of the qualifier than of the head. Finally, over and above this asymmetry there was a further asymmetry in that one of the concepts in each pair tended to have greater weight in the regression than the other – an effect referred to as concept *dominance*. Further research by and colleagues (Storms et al., 1996) suggests that dominance may relate to category size, with smaller categories being more dominant. Another possibility is that dominance relates to some notion of category *coherence*. For example, Patalano, Chin-Parker and Ross (2006) reported that in conjunctions of social categories, people consider that a member of the conjunction will inherit any arbitrary property from a more coherent category in preference to a less coherent category.

The other line of research on conjunctions considered how concept *intensions* are combined. People were asked to offer descriptions of the conceptual "contents" of individual concepts like SPORTS or GAMES, and then to describe the properties of the conjunction 'Sports that are also games'. To improve reliability, a second set of participants were then given a list of all the properties generated to any of the concepts. Different groups judged the extent to which each property could be considered true of each constituent or of the conjunction. In this way the pattern of *attribute inheritance* was explored, using an analogy of the conjunction as the offspring of the two constituent parents.

Unlike the extensional task, there are generally no hard logical constraints that can be "broken" in the attribute judgments. The only hard constraint arises when people consider that a feature or property is a *necessary* attribute of a category, in which case it should also be necessary for conjunctively defined subsets of the category. The results (Hampton, 1987) suggested that people are in fact aware of this general constraint. The rating scale asked for judgments of "how important" an attribute is for a concept, with the top of the scale labeled "Necessary". When an attribute was given a "necessary" rating for one or more constituents, then it also tended generally to be given a necessary rating for the conjunction. A similar constraint was seen at the bottom of the scale, where an attribute could be judged "impossible" for a concept. Attributes that were impossible for a constituent also tended to be impossible for conjunctions (the one exception being pet birds, where it was judged impossible for birds to talk, but not impossible for pet birds to talk).

The detailed story of how attribute inheritance works requires a fairly lengthy description. At a first level, regression equations suggested that importance of an attribute for the conjunction could be approximately quite accurately in terms of a weighted average of its importance for each constituent. Furthermore the same categories that were dominant for determining membership were also dominant for determining attribute importance (hence once again implicating the role of conceptual coherence in the process). There were exceptions to this general rule, including the case of impossible and necessary features already mentioned. The two main sources of exceptions to simple averaging of attribute importance came from the retrieval of facts about the world from long term memory, and the detection of inconsistencies among inherited attributes. Since both of these exceptions require access to knowledge that is *external* to the concepts being combined, they can be termed non-compositional.

The first of these, dubbed *extensional feedback*, showed up in the presence of “emergent attributes” in the conjunction which were not present in either constituent concept. For example pet birds are judged to live in cages, but neither pets nor birds typically do this. The attribute is clearly made available by people recalling visits to pet shops or to friends’ homes where they have seen pet birds. (Alternatively people might use their “background theories” to guess that if a bird was not kept in a cage, it would not remain a pet for very long.) Although the occurrence of emergent features complicates the process of exhaustively predicting the prototypes corresponding to complex expressions – since exhaustively detailing what a subject has received extensional feedback from is practically impossible – it does not prevent the development and testing of the model against empirical data.

The second source of variation from simple attribute inheritance came from conflicting attributes. Pets are warm and cuddly, while fish are cold and slimy. It is not possible for something to be simultaneously warm and cold, and it seems hard to imagine cuddling a cold slimy fish. The result is that one or other of the features has to be dropped – attribute inheritance fails on this occasion. Hampton (1987) found evidence that (at least in the case of Pet Birds) the choice of which feature to drop is determined by trying to maximize coherence (reduce conflicts) through the minimum amount of change (see also Thagard, 1983).

We have summarized research into concept conjunctions quite briefly. There is in addition evidence that conjunctive concept judgments may show compensation with the typicality of membership in one category influencing the degree to which a borderline member of the other category belongs in the conjunction (Hampton, 1996). There is also evidence (Hampton, 1997a) that the effects observed with conjunctions carry over to negated conjunctions (Sports that are not Games) with increased effect size. Work has also been done on disjunctive combinations (Hampton, 1988a), on attribute inheritance with negation (Hampton, 1997a) and on emergent attributes in a wider range of concept conjunctions (Hampton, 1997b). In addition to the forms of non-logical thinking that have been described here many other forms of reasoning deviating from logical norms have been reported, including: the Conjunction Fallacy (Tversky and Kahneman, 1983), the Inverse Conjunction Fallacy (Jönsson and Hampton, 2006), the inclusion Fallacy (Shaffir, Smith & Osherson, 1990), the Premise Specificity Effect and the Inclusion Similarity Effect (Sloman 1993). As with the effects described in this section all of these effects are most naturally explained by reference to prototype-like representations and the operation of similarity-based reasoning.

IV. The Non-Compositionality of the Intensional Route

We noted earlier that since the treatment of conceptual conjunctions within prototype theory deviates from the norms of classical logic it has been held to be an incomplete or even erroneous theory of concepts (Osherson & Smith, 1997). This purported incompleteness has been used to promote frameworks where concepts are understood to be something quite different from a prototype. For instance, Fodor (1998) and Connolly et al. (2007) favour a model – the classical model[13] – where concepts are atomic, i.e. have no internal structure. On this view, conceptual combination is a process where "concepts remain inert under

[13] We will call it the classical model, although as detailed in Smith and Medin's (1981) seminal work, the classical model proposed that concepts have defining features. In contrast Fodor's account eschews definitions, and simply proposes that concepts as they play a role in mental life are unanalyzed atomic symbols.

combination" and "all you get from your concepts and combinatorics is output denoting relations among sets, properties, or individuals (depending on the ontology assumed)" (Connolly et al. 2007, pp. 2-4). So in contrast with prototype theory, the classical model does not hold that the result of combining two concepts is some amalgamate of their prototypes. Instead, the result of combination is a simple structure having the original concepts as proper parts. This difference can be illustrated in the following way for a conjunctive combination ('sports that are also games' for instance).

The classical model: $A_1, A_2 \rightarrow A_1 \wedge A_2$

Prototype Theory: $P_1, P_2 \rightarrow P_3$

As we have seen above, the ways in which prototypes combine via the Intensional Route is not such that only the things falling under P_1 and P_2 fall under P_3. This constraint is however guaranteed on the classical model, by the definition of '$\wedge$'. By being able to account for our conviction that only, say, things that are sports *and* that are games are sports that are also games, the classical model seems to be able to do something which prototype theory is unable to do. It is thus concluded that prototypes cannot be concepts, or at least that two different kinds of concepts need to be invoked. But there might be good reasons to resist this conclusion.

When expressions have a syntactic form that makes their logical form explicit, people are often ready to assent to the validity of simple arguments on the basis of form alone. Thus *modus ponens* arguments such as (1) may be accepted as valid without any information regarding the meaning of the unfamiliar terms

(1) If all lokies pling then all maks flug.

All lokies do pling,

Therefore all maks must flug.

Similarly, people may assent to arguments involving the simplification of conjunction in the same way.

(2) *x* is a flug which is also a loki

Therefore *x* is a flug

Hence, the intuition that ‘x is a sport which is also a game’ entails ‘x is a sport’ might be judged valid on basis of its form alone. Violation of form would explain why we feel that there is something wrong with saying, for instance, that chess might be a sport which is also a game, but that it is not a sport. This formal explanation of our logical intuitions can then be set alongside the alternative explanation of why people are in fact prepared to give inconsistent judgments when form is over-ridden by content.

Another, *semantic* explanation of our logical intuitions is in terms of logical composition rules for prototypes. There is nothing about prototypes that prevent them from being just as amenable to logical combination as are atoms. It was the kind of composition rules used rather than the fact that concepts were treated as atoms that was key to the explanation of our logical intuitions. Hence, prototype theory could match the explanatory ability of the classical model with respect to logical intuitions by borrowing its composition rules without thereby admitting atomistic concepts. On this view the concept corresponding to ‘sports that are also games’ might be derived in two different ways.

$$P_1, P_2 \rightarrow P_1 \wedge P_2$$

$$P_1, P_2 \rightarrow P_3$$

In this way we again get an explanation of why certain categorization decisions that we make are felt to be inconsistent since something falling under P_3 is not equivalent to it falling under P_1 *and* P_2. If this is correct, then even though we now have two kinds of composition rule, there is no reason to multiply kinds of concepts. With the application of one kind of rule the prototypes mesh – and non-logical reasoning results; with the application of the other kind they remain inert – and logical reasoning results. But in neither case did we need to make reference to any other form of concept than a prototype.

If this last proposal is correct then there is at least one way that prototypes can combine that is in accordance with compositionality.[14] However, if compositionality requires that the content of a complex concept is determined by the constituent contents plus the syntax of the combination, *and nothing else*, then the majority of prototype combinations are clearly not compositional. The occurrence of emergent attributes and of inheritance failures through conflict resolution requires that a much wider range of knowledge and belief is brought to bear on the process of understanding a conjunctive concept. This extra input clearly breaks the constraints of compositionality.

The general non-compositionality of prototypes has been a source of philosophical criticism and we take this opportunity to address some of these broader critiques. Fodor and

[14] See Horwich (1997; 2001) for a good discussion of the possibility of dropping ‘the uniformity assumption’, i.e. the assumption that if something is/engenders the meanings of simple expressions, the same kind of thing must be/engender the meaning of complex expressions.

Lepore (1996), for instance, argue that concepts cannot be prototypes since, supposedly, many complex concepts don't have prototypes.

> "However, although 'isn't a cat' [has a definite semantic interpretation] it pretty clearly has no stereotype; and nor do indefinitely many other Boolean complex concepts. There isn't any stereotypic nonprime number, and there isn't anything that is stereotypically pink if it's square" (Fodor and Lepore, 1996, pp. 260-261)

The claim that the listed concepts do not have a prototype is true on an *extensional* understanding of 'prototype'. They clearly don't have instances such that they could serve as 'centers' for their categories. But the common conception of a prototype which we have been assuming in this paper does not equate a prototype with such an instance. Rather, a prototype is understood intensionally in terms of a structured set of weighted features possibly embedded in a causal schema.

There are several possible ripostes to the "concepts without prototypes" argument. First, psychologists have never claimed that ALL concepts are represented in the mind as prototypes. Mathematical concepts such as *i*, the "imaginary" root of -1, are clearly not represented as prototypes, and no-one has ever claimed that they are. Second, it has been shown that while some familiar concepts have strong single prototypes, others are better thought of as being represented by collections of prototypes at a more specific level (Smits et al., 2002). The typical human face for example is probably represented by separate prototypes of a male and a female face, rather than a single sexless prototype. Such issues require empirical verification. So the claim made by prototype theory is not that all concepts must have prototypes, or that a single prototype is all that is needed to represent all concepts.

The situation is more complex than that.

A third possibility that has been suggested in the literature, is that, for instance, 'cat' and 'not a cat' invoke the same prototype but that the similarity measure is used in a different way when deriving typicality judgments for the two categories (not being a cat is then a matter of being sufficiently *dissimilar* to the cat prototype).[15] There is little empirical work on this, but it is plausible to suppose that a theorem would be considered a much better example of not being a cat than would a dog. Yet without similarity-based reasoning and prototypes, there is of course no way to account for such intuitions.

Finally, one can point out the close correspondence between the existence of a prototype and the naturalness of a concept (Osherson). Fodorian examples such as "pink if it's square" are allegedly concepts without prototypes. But equally they are concepts with no likelihood of gaining ground in our conceptual repertoire. What is more, if they did turn out to be valuable for conceptual thinking or communicative purposes, one can suppose that the Intensional route would kick in, and additional inferences and prototypical features would start to accumulate. Conditionals are not ruled out from prototype concepts – "dangerous if male" might apply to a range of mammals, (just as "dangerous if female" might apply to spiders). Negation is not ruled out either. Dead is "not alive" and impossible means "not possible". But most of our negative concepts relate to binary oppositions, so that being dead or being impossible have as many typical entailments as do their opposites.

The point is that Boolean expressions can be as meaningless as you like, and as long as they remain meaningless they are prototype-less. Add in a meaningful context and a prototype will appear. "Not a cat" is not meaningful out of context, but in a given context (e.g. Christmas presents, pets, road kill) it brings entailments that go beyond the simple

[15] See, for instance, Kamp and Partee (1995) and Prinz (2002).

Boolean expression. As mentioned in Section III, Hampton (1997a) investigated prototypes for expressions such as Building that is not a dwelling, or Vehicle that is not a machine, and revealed some of the complex but systematic effects of how negation operates on prototypes within a given context.

Fodor and Lepore have also advanced variants of *the guppy problem* – the problem which we have already shown faced the extensional route – against prototype theory in general.

> *"The problem here is...that…an object's similarity to the prototype for a complex concept seems not to vary systematically as a function of its similarity to the prototypes of the constituent concepts. So, for example, a goldfish is a poorish example of a fish, and a poorish example of a pet, but it's quite a good example of a pet fish."*
>
> *(Fodor and Lepore, 1996, p. 262)* [16]

The point seems to be that what constitutes a prototype for a complex expression depends on what we believe about the world. If we believe that all brown cows are dangerous the feature 'is dangerous' is going to receive a high weight in the prototype for 'brown cow' regardless of the weights the features corresponding to "brown' and 'cow' receive. However, it is not clear why this dependency is supposed to be a problem. Once it is recognized that the derivation of prototypes for complex expressions via the intensional route depends on general knowledge the guppy 'problem' is simply a statement of this dependency. And many theories

[16] This seems to be more or less the same problem that Fodor and Lepore (1991) have advanced elsewhere against variants of inferential role semantics.

of the composition of prototypes have been constructed in full awareness of this. The concept specialization model (Murphy 1988) and the composite prototype model (Hampton 1988) both make explicit reference to general knowledge. Naturally, too heavy an emphasis on general knowledge could render a psychological account vacuous since our minimal understanding of general cognition would mean that predictions would no longer be available. But as has been illustrated repeatedly (and exemplified above) this is not the case with prototype theory. It is in virtue of its predictions concerning categorization, typicality effects and non-logical reasoning that it remains an attractive model. In addition to the examples already cited, further regularities were revealed by Jönsson and Hampton (2007) pertaining to modifier-noun constructions. One of their results concerned sequences of sentences with increasingly modified head nouns such as 'Ravens are black', 'Young ravens are black' and 'Young jungle ravens are black', or 'Sofas have backrests', 'Uncomfortable sofas have backrests' and 'Uncomfortable handmade sofas have backrests'. Replicating an effect reported by Connolly et al. (2007) they observed that the judged likelihood of such sentences decreases with the increase of the number of atypical modifiers. However, they also found significant correlations between the relative likelihoods of other properties between modified and unmodified subjects, indicating that the strength of a feature for the unmodified noun prototype was predictive of its strength for the modified noun phrase concepts. If property A is more true of a concept than property B, then it turns out that A is also more true of the modified concept that is B. Jönsson and Hampton took this as evidence that information from the concept prototype is imported directly into the modified concept, as would be expected from the operation of the Intensional route. [17]

[17] See Jönsson and Hampton (2008) for other remarks pertaining to the effect first reported by Connolly et al. (2007)

Finally, we consider whether the desiderata of a compositional theory could equally be provided for by prototype theory. The two main planks of the argument for compositionality are *productivity* and *systematicity*. We will consider each in turn.

To account for productivity – how a system of representation can contain ‘an infinite amount of syntactically and semantically distinct symbols’ (Fodor and Lepore 2002, p.1) – a theory needs to specify recursive rules for combining concepts. It is the recursion that is critical for productivity, rather than the nature of the concepts. Most of the intensional prototype models do embody recursive rules; the composite prototype model for instance can be applied recursively to generate a concept such as “Recreations that are games, hobbies, sports and dangerous activities”. Wisniewski’s (1997) model for noun-noun combinations can generate chains of concepts such as “Apartment dog” “Apartment dog coat” “Apartment dog coat cleaner appointment cancellation notification ticket stub residue” and so forth. So there is no barrier to prototype combinations being productive.

A conceptual system being ‘systematic’ amounts to it containing “such families of syntactically and semantically related but distinct expressions such as ‘John loves Mary’, ‘Mary loves John’, ‘John loves John’ and so forth” (Fodor and Lepore 2002, p.2). There is again nothing that prevents prototype models of concept combination from capturing this intuition. The composite prototype model, for instance would imply that anyone who can understand “Sports that are also Games” should be able to understand “Games that are also Sports”, or “Sports that are also Sports” and so on. Having a fixed set of syntactic combination rules applied to a finite set of atomic elements is sufficient to account for productivity and systematicity. But it is by no means necessary. Productivity and systematicity are a by-product of a compositional semantics, but they may also be found with non-compositional systems.

Productivity and systematicity do not favor either model of concepts. In fact, it might even be the case that prototype theory has the explanatory upper hand pertaining to these phenomena. It can be remarked, for instance, that systematicity doesn't hold to the same extent for all complex concepts. Where 'Mary loves John' and 'John loves Mary' are both unproblematic it is much easier to understand 'Mary loves Ice-cream' than 'Ice-cream loves Mary'. Given a conception of concepts according to which they have sub-lexical semantic content, the differences can be explained. But on the classical model where concepts are taken to be atomic no such explanation is possible. Since there are many instances where systematicity is constrained in this way (for instance 'brown cow' and 'prime number' are straightforward, 'brown number' and 'prime cow' are less so), the explanatory scope of the classical theory is more limited. The theory states that for any expression such as "Mary loves John" there will be at least some other expression in which each concept plays the same semantic role ("Mary hates Peter", "Joan loves Jill" or "John plays tennis"). But it has nothing to contribute to the question of where the limits of systematicity lie. On the other hand, a satisfactory prototype account of the concept of "love" will provide an answer to just that question. What sorts of things can love, what kinds of love are possible between what sorts of things, and what sorts of things can be loved in what ways are all important parts of the concept, and each in its turn will be subject to typicality effects. Mary is a typical subject for the verb, a plant is a less typical subject ("the dahlia loves a sunny aspect"), while an inanimate artifact is more atypical again ("my car loves the open road").

V Conclusion

We have tried to illustrate the many ways in which a prototype based approach to conceptual combination, while non-compositional, has yet been successful in providing explanations and predictions pertaining to our use of language and to our conceptual competence. A wealth of phenomena pertaining to typicality effects, vagueness and instances of non-logical reasoning can be explained given that we adopt an account based on the basic tenets of prototype theory.

References:

Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: from feature lists to recursive frames. In I.van Mechelen, J. A. Hampton, R. S. Michalski, & P. Theuns (Eds.), *Categories and Concepts: Theoretical Views and Inductive Data Analysis* (pp. 97-144). London: Academic Press.

Billings, R. S., & Marcus, S. A. (1983). Measures of compensatory and noncompensatory models of decision behavior: process tracing versus policy capturing. *Organizational Behavior and Human Performance, 31,* 331-352.

Chater, N., & Oaksford, M. (1999) Ten years of the rational analysis of cognition. *Trends in Cognitive Science, 3,* 57-65.

Cohen, B., & Murphy, G. L. (1984). Models of Concepts. *Cognitive Science*, *8*, 27-58.

Connolly, A. C., Fodor, J. A., Gleitman, L.R., & Gleitman H. (2007). Why Stereotypes Don't even Make Good Defaults. *Cognition, 103,* 1-22.

Dennis, I., Hampton, J. A., & Lea, S. E. G. (1973). New Problem in Concept Formation. *Nature, 243,* 101-102.

Evans, J. St. B. T. (1982). *The Psychology of Deductive Reasoning*. London: Routledge Kegan & Paul.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and Reasoning.* Hove: Psychology Press.

Fodor, J. A. (1998). *Concepts*: *Where Cognitive Science Went Wrong.* Cambridge, MA: MIT Press.

Fodor, J. A., & Lepore, E. (1991). Why Meaning (Probably) Isn't Conceptual Role. *Mind and Language, 6,* 329-343.

Fodor, J. A., & Lepore, E. (1996). The Red Herring and the Pet Fish: Why Concepts Still Can't Be Prototypes. *Cognition, 58*, 243-276.

Fodor, J. A., & Lepore, E. (2002). *The Compositionality Papers.* Oxford: Oxford University Press.

Gigerenzer, G., Todd, P., & the ABC Research Group (1999). *Simple Heuristics that Make Us Smart*. Oxford: Oxford University Press.

Goldstein, R. (1983). *The Mind-Body Problem*. New York: Penguin Books.

Hampton, J. A. (1979). Polymorphous Concepts in Semantic Memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-461.

Hampton, J. A. (1987). Inheritance of Attributes in Natural Concept Conjunctions. *Memory & Cognition, 15*, 55-71.

Hampton, J. A. (1988a). Disjunction of Natural Concepts. *Memory & Cognition, 16,* 579-591.

Hampton, J. A. (1988b). Overextension of Conjunctive Concepts: Evidence for a Unitary Model of Concept Typicality and Class Inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 12-32.

Hampton, J. A. (1991). *The Combination of Prototype Concepts*. In P. J. Schwanenflugel (Ed.), *The Psychology of Word Meanings*, (pp. 91-116). Hillsdale, Lawrence Erlbaum Associates.

Hampton, J. A. (1995). Testing Prototype Theory of Concepts. *Journal of Memory and Language, 34,* 686-708.

Hampton, J. A. (1996). Conjunctions of Visually Based Categories: Overextension and Compensation. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 22, 378-396.

Hampton, J. A. (1997a). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition, 25*, 888-909.

Hampton, J. A. (1997b). Emergent attributes in conceptual combinations. In T.B.Ward, S. M. Smith, & J. Viad (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes* (pp. 83-110). Washington DC: American Psychological Association Press.

Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition, 65*, 137-165.

Hampton, J. A. (2006). Concepts as Prototypes. In B.H.Ross (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 46* (pp. 79-113). Amsterdam: Elsevier.

Hampton, J. A. (2007). Typicality, Graded Membership and Vagueness. *Cognitive Science, 31*, 355-383.

Horwich, P. (1997). The Composition of Meanings. *Philosophical Review, 106*, 503-532.

Horwich, P. (2001). Deflating Compositionality. *Ratio, 14*, 369-385.

Jönsson, M. L. (2008). Compositionality: The Structural Path to Meaning. Unpublished MS, Department of Philosophy, Lund University.

Jönsson, M. L., & Hampton, J. A. (2006). The Inverse Conjunction Fallacy. *Journal of Memory and Language*, *55*, 317-334.

Jönsson, M. L., & Hampton, J. A. (2007). The modifier effect in within-category induction: Default inheritance in complex noun phrases. Unpublished MS: Department of Psychology, City University, London.

Jönsson, M. L., & Hampton, J. A. (2008). On prototypes as defaults. (Comment on Connolly, Fodor, Gleitman and Gleitman, 2006). *Cognition*, *106*, 913-923.

Kamp, H., & Partee, B. (1995). Prototype Theory and Compositionality. *Cognition, 57*, 129 - 191.

Keefe, R. & Smith, P. (1997). Theories of vagueness. In R. Keefe & P. Smith (Eds.), *Vagueness: A Reader* (pp. 1-57). Cambridge: MIT Press.

Lahav, R. (1989). Against Compositionality: The Case of Adjectives. *Philosophical Studies, 57*, 261-279.

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12,* 529-562.

Osherson, D. N. (1978). Three conditions on conceptual naturalness. *Cognition, 6,* 263-289.

Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition, 11,* 35-58.

Osherson, D. N., & Smith, E. E. (1982). Gradedness and conceptual conjunction. *Cognition, 12,* 299-318.

Osherson, D. N., & Smith, E. E. (1997). On typicality and vagueness. *Cognition, 64,* 189-206.

Patalano, A. L., Chin-Parker, S., & Ross, B. H. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language, 54*, 407-424.

Pelletier, F. J. (1994). Semantic Compositionality: The Argument from Synonymy. In R. Casati, B. Smith, & G. White (Eds.) *Philosophy and the Cognitive Sciences* (pp. 311-317). Vienna: Höldar-Pichler-Tempsky.

Pelletier, F. J. (2000). Semantic Compositionality: Free Algebras and the Argument from

Ambiguity. In M. Feller, S. Kaufmann, & M. Pauly (Eds.) *Formalizing the Dynamics of Information* (pp. 207-218). Stanford: CSLI Publications:

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353-363.

Prinz, J. J. (2002). *Furnishing the Mind.* Cambridge, MA: MIT Press.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14,* 665-681.

Rosch, E. R., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Rosch, E. R. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-232.

Schiffer, S. R. (1987). *The Remnants of Meaning*. Cambridge, MA: The MIT Press.

Schiffer, S. R. (2003). *The Things we Mean*. Oxford: Oxford University Press.

Shaffir, E., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. *Memory & Cognition, 18*, 229-239.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, *25*, 231-280.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, *8,* 337-361.

Smits, T., Storms, G., Rosseel, Y., & De Boeck, P. (2002). Fruits and vegetables categorized: An application of the generalized context model. *Psychonomic Bulletin and Review, 9*, 836-844.

Stigler, S. M. (1999). *Statistics on the Table.* Cambridge MA: Harvard University Press.

Storms, G., De Boeck, P., Hampton, J. A., & Van Mechelen, I. (1999). Predicting conjunction typicalities by component typicalities. *Psychonomic Bulletin and Review, 4*, 677-684.

Storms, G., De Boeck, P., van Mechelen, I., & Ruts, W. (1996). The Dominance Effect in Concept Conjunctions: Generality and Interaction Aspects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1-15.

Storms, G., De Boeck, P., van Mechelen, I., & Ruts, W. (2005). Not guppies, nor goldfish, but tumble dryers, Noriega, Jesse Jackson, panties, car crashes, bird books, and Stevie Wonder. *Memory & Cognition, 26,* 143-145.

Storms, G., Ruts, W., & Vandenbroucke, A. (1998). Dominance, overextensions, and the conjunction effect in different syntactic phrasings of concept conjunctions. *European Journal of Cognitive Psychology, 10*, 337-372.

Taylor, J. R. (2003). *Linguistic Categorization (3rd ed.)*. Oxford: Oxford University Press.

Thagard, P. (1983). Conceptual combination: a frame-based theory. Paper presented to the Annual Meeting of the Society for Philosophy and Psychology, Wellesley, June.

Tversky, A. and D. Kahneman (1983). Extensional versus intuitive resoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293-315.

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, *4,* 167-183.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control, 8,* 338-353.

Figure 1

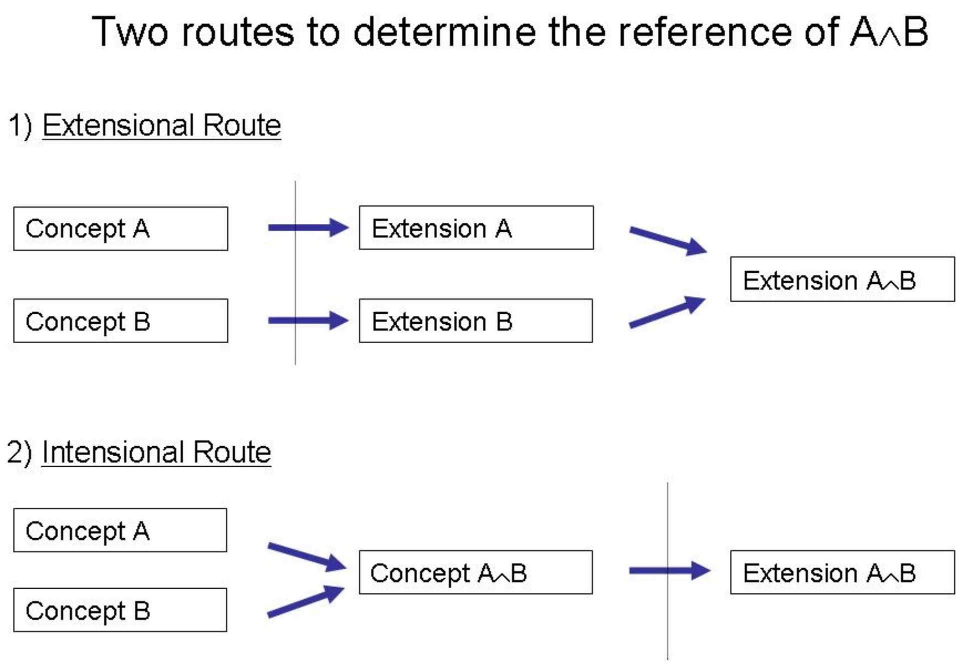

Two routes to determine the reference of A∧B
1) Extensional Route
Concept A
Concept B
Extension A
Extension B
Extension A∧B
2) Intensional Route
Concept A
Concept B
Concept A∧B
Extension A∧B

Figure 2

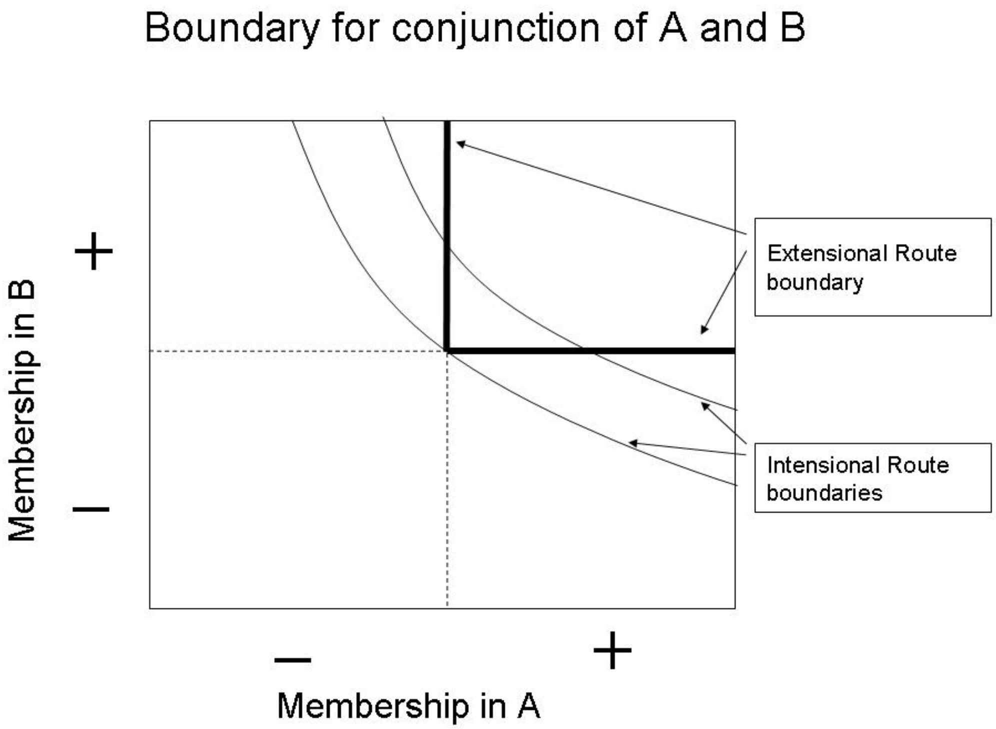
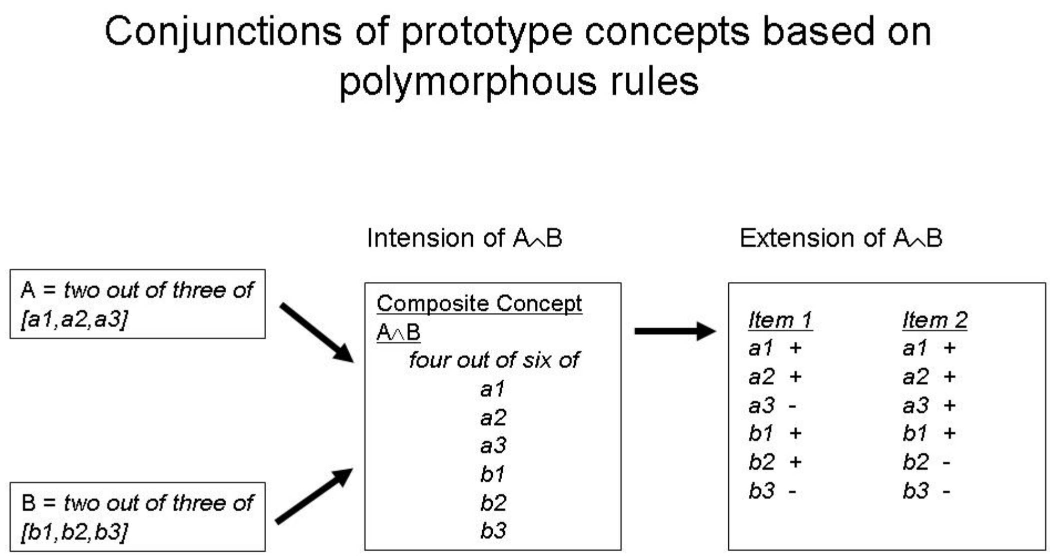
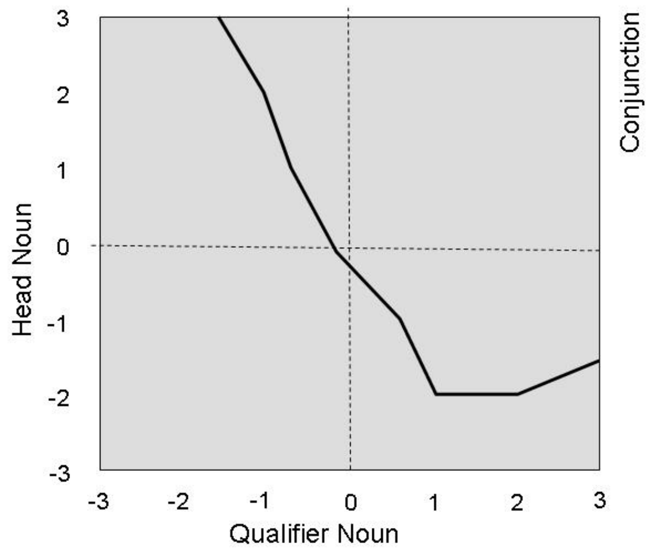

Boundary for conjunction of A and B
Extensional Route boundary
Intensional Route boundaries
Membership in B
+
–
–
+
Membership in A

Figure 3



The composite concept is generated by pooling the features of each concept and requiring a similar level of feature match (66% match). Whereas Item 1 correctly belongs in A and B and in A∧B, Item 2 belongs to A∧B but *not* to B. Weak membership in B has been compensated for by high typicality in A.

Figure 4



Overextension pattern seen in Hampton (1988). The line drawn on the figure is the boundary for the conjunction.